

# TESTING IN CORE COURSES

- I. [INTRODUCTION](#)
- II. [SUCCESSFUL COMPLETION OF CORE TRAINING - WHAT DOES IT MEAN?](#)
  - [Selection of Employees](#)
  - [Retain or Terminate?](#)
- III. [MINIMUM STANDARDS FOR SELECTION, TITLE 15](#)
  - [Section 131](#)
- IV. [TYPES OF TESTS IN CORE TRAINING](#)
  - [Job Knowledge Tests](#)
  - [Who Should Write the Multiple Choice Test Questions?](#)
  - [Written Skills Test](#)
  - [Behavior Skills Tests](#)
- V. [ADMINISTRATION OF THE TESTS](#)
- VI. [SCORING THE TESTS](#)
  - [Setting the Cut-Off Score for Job Knowledge Tests](#)
  - [Scoring Written Skills Test](#)
  - [Behavior Skills Test](#)
  - [Remediation/Failing Tests in Core Courses](#)
- VII. [PHYSICAL SKILLS TESTING \(An Interview from the Standard, 1987\)](#)
- VIII. [JOB KNOWLEDGE TEST ITEM WRITING - MAJOR PRINCIPLES](#)
- IX. [ANATOMY OF A TEST ITEM](#)
- X. [TEN RULES FOR TEST ITEM WRITING \(with examples\)](#)
- XI. [TEN RULES CHECKLIST](#)
- XII. [STEPS TO WRITING TEST QUESTIONS](#)

## I. INTRODUCTION

As a core provider you will be required to test the trainees on mastery of the subject matter during the course presentation. A specified number of hours is built into the overall core hours total to accommodate the testing process. Each core course has a different number of hours allocated for testing. Please see the core training manual to determine this number.

It is important to view tests in core as part of the overall selection process. A trainee's failure may mean the loss of employment. This means the tests must be constructed, administered, scored and reported appropriately. This publication is intended to help providers and instructors fulfill this testing responsibility

## II. **SUCCESSFUL COMPLETION OF CORE TRAINING - WHAT DOES THIS MEAN?**

The Board of Corrections regulations pertaining to the selection of local corrections personnel specify as one of seven selection criteria **"successful completion of core training."** (Title 15, Section 131, subsection a, 4.) As part of this standard, an individual must demonstrate mastery of job related material by passing several tests embedded in the core training course. Passing these tests is a continuation of the selection process: failing these tests could result in the employee's termination.

It is important to the trainee and the agency that the core tests make an accurate assessment of the trainee's mastery. Because of this, care is required when developing the tests used in the core training course. The tests must meet the same validity and fairness guidelines as other steps in the selection process (entrance exam, oral interview, background investigation, psychological, etc.).

### **Selection of Employees**

Employee selection criteria generally fall into 2 categories: before-hire and after-hire characteristics. The before-hire characteristics are the knowledge, skills and abilities an applicant brings to the job, such as the ability to read and write, follow directions and communicate orally, be psychologically and medically fit. After-hire characteristics include the ability to demonstrate mastery of knowledge and skills taught in training. These include job knowledge, physical performance skills and application of job specific information.

Assessment of a candidate's performance in training can be an excellent way to identify who will be successful on the job and who will fail; but, the assessment has to be valid and fair.

### **Retain or terminate?**

When an agency administrator receives information that an entry-level employee has not successfully completed the core training curriculum, s/he must make a hard choice. If the agency keeps the individual employed, the agency may be liable for negligent retention. If the agency terminates the employee, the agency may be liable for negligent termination. The validity and fairness of the tests administered in core training will be scrutinized in either case.

Given these possible consequences, what expectations should an agency have of the core training delivery system, (be it in-house or through an outside provider)? Is the agency expecting the provider to "wash out" an employee who fails the core tests? Or, is the agency expecting significant remediation until mastery is demonstrated? And, what happens if, after remediation, the employee still is not able to pass?

**Note:** STC requires that scores on all core tests be provided to the agency within 30 days of course completion so that employing agencies can evaluate the suitability of their employees for permanent hire.

### **III. MINIMUM STANDARDS FOR SELECTION, TITLE 15**

#### **Section 131.**

- (a) In addition to the requirements set forth in Section 830 et seq. of the Penal Code and Section 1029 et seq. of the Government Code, the standards set forth below shall apply. These standards for entry probation officer positions, entry juvenile counselor positions, and entry corrections officer positions shall include but not be limited to the following:
  - 1. Basic abilities and other characteristics important for successful job performance as demonstrated by passing the Board's written examination. An alternative examination may be substituted pursuant to Section 132.
  - 2. Competence in oral communication as demonstrated in an interview.
  - 3. Past behavior compatible to job requirements as demonstrated by a background investigation.
  - 4. Competence in the knowledge, skills and abilities necessary for entry-level job performance, as demonstrated by successful completion of the required core curriculum.
  - 5. Competence in the performance of entry-level duties as demonstrated by successful completion of the probationary period.
  - 6. The ability to perform the essential job functions of the position as demonstrated by meeting the Board's current guidelines for Vision, Hearing, and Medical Screening.
  - 7. A minimum of 18 years of age prior to appointment.
- (b) The level of competence in (a) (2) and (a) (3) above shall be commensurate with the needs of the individual job classifications of each county or city. The level of competence in (a) (1) above shall be commensurate to the cutoff score that is chosen by the county or city and is consistent with research validation.

#### **IV. TYPES OF TESTS IN CORE TRAINING**

**There are three specific types of tests in the core courses:**

- Job Knowledge Test (JKT)
- Written Skills Test (WST)
- Behavior Skills Test (BST)

Job Knowledge Tests are paper and pencil tests that are usually multiple choice. Written Skills Test require the trainee to write a simulated job document such as a portion of a report, interview questions, or incident report. Behavior Skills Tests require the trainee to physically demonstrate a job skill such as donning a Self-Contained Breathing Apparatus or dragging a weighted dummy a certain distance within a prescribed amount of time.

The type of test required is indicated next to each performance objective in the core training manual. The notation will specify which type of test (JKT, WST or BST). If the notation is JKT, the performance objective will also indicate how many questions are to be asked for that performance objective.

For instance, the core training manual will read as follows:

Example: Performance Objective 5.5.5 List common hiding places for contraband.JKT: **1** item.

##### **Job Knowledge Tests**

If the type of test indicated is a Job Knowledge Test, the number of test questions is the number in bold above. Providers must test for the minimum number of questions specified. Providers are encouraged to allow trainees one-half minute per test question.

##### **Who Should Write the Multiple Choice Test Questions?**

In most case, the instructors should write the test questions because they will be able to link the questions most accurately to the lesson plans. It is also recommended that instructors write extra test questions to develop a bank of test questions to draw on:

- in case of remediation,
- as substitutes for poorly written questions, and
- to alternate questions for test security

The Board of Corrections does not maintain a test item data bank but is available for assistance in constructing test questions.

Since there is no practical way of measuring every single aspect or detail of the job through a testing process tests in core training measure mastery by sampling the desired level of competence. Therefore, constructing a representative and meaningful sample of the important aspects of training is a major part of test construction.

##### **Written Skills Test**

The procedure for Written Skills Testing is essentially laid out in the performance objective. For example:

Performance Objective 11.1.4: Given a description of an incident, generate an appropriate report, using the report writing manual, department policy manual and all relevant codes as reference. WST #6

This means that the Written Skills Test (labeled WST #6) is the trainee's preparation of the report. The report is to be written during the classroom instruction time and graded by the instructor during non-classroom hours.

## **Behavior Skills Tests**

The procedure for Behavior Skills Testing is also specified in the performance objective. For example:

Performance Objective 21.1.4: Demonstrate ability to quickly drag a 165 lb. bag or dummy by the handle or strap for a distance of 20 feet or more over a carpeted or flat, mowed grass surface. BST #10.

This means that the successful accomplishment of this performance objective constitutes the Behavior Skills Test.

## **V. ADMINISTRATION OF THE TESTS**

In addition to sound test construction techniques, there are several other issues that relate to core tests. As part of the training delivery process, agencies and providers are encouraged to examine the following areas of core test administration:

1. Who should develop the tests? Should this be the instructor or someone skilled at test item writing?
2. What's wrong with "teaching to the test"?
3. After the test, should instructors go over each test question with the class so the trainees know what they got wrong or right?
4. Are test pass points (cutoff scores) set so that there is a reasonable expectation that those who pass the test have mastered the material?
5. How are trainee challenges to the test handled?
6. What is the test retake policy? Do the trainees take the same test over again or is there a bank of test questions available to sample the trainee's mastery?
7. Is there a written policy about how to handle trainees who fail? Does the policy appear fair and is it applied evenly?
8. Under what conditions are the trainees tested? Is there anything going on during testing that would distract the trainee or negatively impact their performance?
9. What kind of documentation is maintained that links the test to what is taught? Do the tests emanate from the lesson plans?
10. How is test security handled? Are the tests proctored? Where are the tests and test answers kept?

## VI. SCORING THE TEST

**General Principle:** The provider is responsible for determining if the trainee has demonstrated entry-level mastery of the curriculum.

### Setting the cut-off score for Job Knowledge Tests

Providers must determine what constitutes a passing score on the job knowledge tests. The cut-off point cannot be determined by the Board of Corrections because the test questions are generated by the instructor's lesson plan. There may be wide variation in the difficulty, number and emphasis of the questions throughout the state. The only meaningful way to accommodate these variations is to allow the provider to determine what constitutes a passing score.

The provider's cut-off score should take into consideration the following item analysis of each question:

- The level of difficulty of the question,
- The average response rate, over several deliveries, for the question,
- The number of people who answered the question correctly,
- The correlation between performance on other test questions and the question being analyzed – in other words, did someone who scored well on most other questions fail this question?
- The number of questions the provider decides to ask (above the minimum required) for each objective,
- Compelling feedback from trainees regarding ambiguity of the question.

After such an analysis of test items, the provider makes the determination as to whether the trainee mastered the material.

Typically, providers use a percentage cut-off such 70% or 80%. Another commonly used scoring report form is raw scores, such as 7 out of 10 correct.

### Scoring Written Skills Tests

These tests are scored as pass/fail only. The key to evaluating mastery of the curriculum by the Written Skills Tests lies in the specifications of the performance objective. For instance, Performance Objective 9.10.1 in the Juvenile Counselor Core reads:

Given a sample description of various juveniles, generate work assignments for each juvenile according to:

- past behavior
- classification status
- nature of work
- ability to use tools or learn
- maturity level of juvenile

The Specifications for the Written Skills Test are laid out in the performance objective. To score the test results, the instructor should look for the 5 elements listed in the performance objective and score the trainee according to how many of the elements were addressed and how well they were addressed.

In addition, the instructor should also base scoring on how well the trainee incorporated previous elements of the course that build toward the specific objective called for in the Written Skills Test.

The instructor should also review the trainee's grammar and writing skills as well as content. However, this is a secondary consideration in the scoring. If the trainee is in need of improvement in writing skills, the instructor may include this through the Individual Trainee Evaluation Form or other feedback to the agency training manager.

## **Behavior Skills Tests**

These tests are also scored as pass/fails. The specifications of passing these tests are embedded in the performance objective. For instance, performance objective 21.1.4 of the Corrections Core Course requires the trainee to:

Demonstrate the ability to quickly drag a 165 lb. bag or dummy by the handle or strap for a distance of 20 feet or more over a carpeted or flat, mowed grass surface.

In this test, the trainee can either perform the task or not and thus, receives a pass or fail score. Mastery of the curriculum in the Behavior Skills Tests is based on the research conducted through the statewide job analysis.

## **Remediation/Failing Tests in Core Courses**

Core training providers should have in place prior to the first course delivery a written policy on how they will handle a test failure. This policy should be followed consistently for all trainees. For instance, if one re-take is available to a trainee who fails a test, then one re-take should be available for all trainees. Re-takes for Job Knowledge Tests should be made using different test questions than the original test.

If a trainee is unable to pass a test after the remediation and re-take is performed; the training provider should be in contact with the participating department. It may be unwise to dismiss a trainee from the rest of the core course just because they are unable to pass a portion of it. It may be that the participating department wishes to remediate the trainee at a later date and re-test the employee on-site.



## **VII. PHYSICAL SKILLS TESTING (AN INTERVIEW FROM THE STANDARD, 1987)**

The terms, concepts, and issues related to physical performance, conditioning and agility are sometimes confusing components of the employee selection process. To help clarify the issues, STC Consultant and Newsletter Editor Leroy Ford interviewed industrial Psychologist John Kohls.

### **The Standard:**

**The Board of Corrections Selection and Training Standards have undergone big changes in the last year. One change bears some close examination: physical performance and agility testing, physical conditioning training and subsequent testing. In order that Board of Corrections requirements will be well understood and well administered, it would be useful to gain some knowledge regarding this topic.**

**Would you explain some of the basics and define for us physical performance testing? First, why we do it and secondly, what are the expected outcomes?**

### **Kohls:**

Everything we do in the area of employee selection and training is based on tasks the job incumbent performs.

When we did the original job analysis we found a large number of important and frequent tasks that were being performed by people in corrections jobs, the juvenile counselor job, and the probation officer job. When we further analyzed some of these tasks we found that some of them had implications for physical performance: that is, in order to perform the tasks successfully, the job incumbent had to demonstrate a certain amount of physical prowess. Each time we find one of those tasks we must address the issue of whether task performance should be measured before someone is hired, before they go into a training program, toward the end of the training program, or back on the job. Regardless of where we decide the testing should occur, our goal would be to demonstrate that the person who is applying for the job can perform tasks that are physically demanding. Therefore, a physical performance test is a test which we give to applicants or students in training to verify the fact that these people can perform the physically demanding work.

### **The Standard:**

**That tells us something about physical performance testing. To make this more clear, we have two alternatives. We can either test in connection with hiring, or we can test at a later time. What are the advantages and disadvantages of each of these?**

### **Kohls:**

If you decide that a physical performance test should be part of the entry-level screening process, it means you have concluded that the person needs to demonstrate that they possess physical abilities prior to entering into a training program. There are some physical abilities that you would not expect a person to acquire during a training program and, therefore, they must possess them before they are hired. For example, there may be a certain level of strength required by certain job tasks, the kind of strength which would be unlikely for a person to develop during a relatively short training program, like six weeks or even two or three months.

### **The Standard:**

**Therefore...**

### **Kohls:**

The decision about whether or not to give the physical performance test as part of the screening process depends partially on whether you think somebody who doesn't have that ability can acquire the ability during training. So, it becomes a judgment call. One advantage for placing the physical performance test in training, as opposed to

including it as part of the entry-level selection process, is that it might be easier to defend. For example, if you have selected a person for a position, put them into training, allow them to practice on tasks, allow them to improve to the extent that they can, and then find that they simply can't perform the task, it is more defensible in that case to reject that person than if you rejected them without their having the opportunity to practice the task and improve their skill.

**The Standard:**

**So, can I assume for most corrections jobs walking, talking, and being able to see would be characteristics that you would want to test prior to a training course? Is that correct?**

**Kohls:**

Assuming that walking, talking, and seeing is required on the job, I would think that you would want to verify that this person possesses these abilities prior to hiring them.

**The Standard:**

**But, certain technical kinds of things - strength and certain other physical skills could be reasonably acquired during a short period of training?**

**Kohls:**

That's right. I think that we know now that certain physical capacities can be acquired relatively quickly if we work on them with a regular routine of exercise.

**The Standard:**

**But, certainly those capacities are limited to those which would be connected with the job, not outside that. We're not talking about running long distances or lifting extremely heavy weights.**

**Kohls:**

The most important consideration is that it be part of the job. However, it is not quite that simple. We always run into a problem of what should be the cutoff score on a physical performance test. If it is a question of running, how fast does a person have to be? Certainly, if you have to chase someone to apprehend them, the best situation would be if the good guys were able to catch world-class sprinting bad guys. However, we simply can't make those kinds of demands.

**The Standard:**

**So, there has to be some reasonableness to the expectations. We can't place the expectations so high that few people could pass them.**

**Kohls:**

The question boils down to, what is your definition of reasonable? There are several. One definition of reasonable is that the cutoff score should not be so restrictive that currently successful job incumbents could not pass the test. In other words some would say that an entry-level person shouldn't have to possess any physical ability that a successful incumbent doesn't have.

Another approach would be to assume that a certain small percentage of the job incumbents are not performing up to par and, therefore, the cutoff score should be set at point which would pass most of the incumbents but not all of them.

The third approach would consist of an attempt to try to assess how much of the ability/capacity would be required in the average incident. For example, when someone has to run quickly over some distance, typically, how quickly do they have to run and how far do they have to run? Using this kind of data one can set a cutoff score that is consistent with a typical occurrence.

**The Standard:**

**A typical occurrence may differ from organization to organization just as physical plants may differ, one from another?**

**Kohls:**

That's true not only of physical characteristics but of all the tasks that are performed by job incumbents. Therefore, it was the approach of the Board to set minimum standards which would apply equally well to all locations in which the job was performed.

**The Standard:**

**Allowing the employing agency to set other standards which might apply specifically to their particular work place?**

**Kohls:**

That is correct. I don't know if it has ever been articulated but I'm sure that the Board would want to encourage local agencies to set their standards as high as they can possibly set them.

**The Standard:**

**Based on the job to be accomplished?**

**Kohls:**

Yes, always with that caveat - that the standards be consistent with the job that is being performed.

John W. Kohls, Ph.D. has been a consultant to the STC Program on Selection and Training Standards since 1981. He is a former Director of the POST Commission's Research and Evaluation Unit.

## **VIII. JOB KNOWLEDGE TEST ITEM WRITING MAJOR PRINCIPLES**

- I. Does it measure something important?
- II. Does it differentiate between those who know the material vs. those who don't?
- III. Does it measure what you want it to measure?
- IV. Is it fair, valid and reliable?
- V. Does it link to the lesson plan? Did you teach what it tests?

## **IX. ANATOMY OF A TEST ITEM**

**Stem:** The question being asked

**Example:** The capitol of California is...

**Answer:** The correct answer

**Example:** Sacramento

**Distractors:** Incorrect answers.

**Example:** Los Angeles, San Francisco, and San Diego

## **X. TEN RULES FOR TEST ITEM WRITING (with Examples)**

**Rule #1:** The stem of the item should be a complete expression of the problem with only the essential information missing.

After reading the stem, the candidate should know exactly what response is expected from the choices offered. The central issue in each item should be clearly and accurately stated. If you can cover up the alternatives with your hand and still understand what problem the item poses, you have satisfied this rule.

Poor Example:

"Sacramento is...."

1. the capitol of California
2. smaller than Los Angeles
3. hot in the summer
4. the city of trees

Better Example:

"The capitol of California is named...."

1. Los Angeles
2. San Jose
3. San Francisco
4. Sacramento

**Rule #2:** As much of the item as possible should be included in the stem.

Including as much information as possible in the stem accomplishes two things. First, the problem is more completely developed in the stem and, therefore, easier to understand. Second, the amount of reading is reduced by eliminating redundant information from the alternatives and placing it in the stem.

Example of a poor item:

The Community Release Board is responsible for

1. supervising probationers
2. supervising state parolees
3. supervising federal parolees
4. supervising state inmates on release time

Example of a better item:

The Community Release Board is responsible for supervising

1. probationers
2. state parolees
3. federal parolees
4. state inmates on release time

**Rule #3:** Avoid negatively worded items whenever possible.

Sometimes it is easier to write an item with three correct alternatives and one wrong one. The stem is then written in the negative so that the wrong alternative is the correct answer. This practice should be avoided whenever possible because it is unnecessarily confusing.

Also, it may condition a trainee to remember only the wrong answer rather than the correct answer. An exception to this rule occurs when the wrong alternative has dire consequences and training has emphasized what not to do rather than what to do.

Example of a poor item:

Which of the following is not a consideration when making an arrest for violation of probation?

1. nature of the violation
2. distance to location for the arrest
3. availability of supervision
4. availability of assistance in making the arrest

Example of a better item:

Which of the following is the most important consideration when making an arrest for violation of probation?

1. time of day
2. availability of assistance in making the arrest
3. nature of the violation
4. availability of supervision

**Rule #4:** All of the alternatives should be grammatically consistent with the stem.

In an incomplete statement each of the alternatives must combine with the stem into a grammatically correct sentence.

Example of a poor item:

A peace officer can hardly wait to get off duty to go to his favorite bar. Since lunch he has been thinking about drinking his customary five or six martinis. The officer can be called an

1. alcoholic
2. social drinker
3. normal drinker
4. plateau drinker

Example of a better item:

A peace officer can hardly wait to get off duty to go to his favorite bar. Since lunch he has been thinking about drinking his customary five or six martinis. The officer can be called

1. an alcoholic.
2. a social drinker
3. a normal drinker
4. a plateau drinker

**Rule #5:** An item should contain only one correct or clearly best answer.

The answer to an item should not be a matter of conjecture among experts. Competent subject matter experts should be able to agree on one best answer; otherwise, the item should be rewritten or discarded.

Example of a poor item:

The best restraint device is the

1. Smith and Wesson handcuff
2. "belly" chain with "irons"
3. leather restraint
4. strait jacket

Example of a better item:

The restraint device with the most durability is the

1. Smith and Wesson handcuff
2. "belly" chain with "irons"
3. leather restraint
4. strait jacket

**Rule #6:** The relative length of the alternatives should not provide a clue to the correct answer.

There is a tendency to make the answer longer than the distractors. This occurs when qualifiers must be added to the answer because the answer is not correct under all circumstances or because the correct answer is unavoidably lengthy.

Example of a poor item:

If an institutional officer suspects an escape-in-progress, he should

1. respond as usual but stay alert
2. secure the perimeter of the area before entering
3. wait 10 minutes before entering the area
4. get sufficient staff support and enter the area from an unexpected direction, maintaining alertness

Example of a better item:

If an institutional officer suspects an escape-in-progress, which of the following techniques would be most appropriate to reduce the danger?

1. respond as usual but stay alert
2. secure the perimeter of the area before entering
3. wait 10 minutes before entering the area
4. enter the area from an unexpected direction

**Rule #7:** Avoid providing a clue to the answer by using the same or similar words in the stem.

Example of a poor item:

During school hours an adult is found loitering in the hallway of a junior high school by a school official. When asked to leave, he refuses. What is the most appropriate charge?

1. prowling
2. trespassing
3. annoying children
4. loitering about a school



Example of a better item:

During gym class a teacher noticed a stranger on the playground watching the children. The teacher asked him to leave, but he refused. What crime did the man commit?

1. prowling
2. trespassing
3. annoying children
4. loitering about a school

**Rule #8:** Avoid the use of all-inclusive terms such as never, always, sole and all in distractors.

The test-wise student will eliminate these alternatives and select the alternative that contains delimiters such as sometimes, generally, and usually.

**Rule #9:** Avoid alternatives such as "none of the above" or "all of the above."

Because it is frequently difficult to write more than one or two good distractors, item writers sometimes resort to special alternatives such as "none of the above" or "all of the above." These alternatives are only appropriate in certain circumstances. For example, "none of the above" can sometimes be used to increase the difficulty of math items. Unless the item writer is well versed in how to use these special alternatives they should be avoided.

An exception to this rule exists for criminal law items where it is important for the student to know when a crime was committed. Under these circumstances, an alternative such as "none of the above" or "no crime was committed" is acceptable. When constructing a test composed of such items, the special alternative should be the answer for approximately 25 percent of the items assuming that the test is composed of four-choice items.

**Rule #10:** The distractors should be plausible enough to attract trainees who do not know the correct answer.

## XI. TEN RULES CHECKLIST

<b>Rule #1:</b>	The stem of the item should be a complete question.
<b>Rule #2:</b>	As much of the item as possible should be included in the stem.
<b>Rule #3:</b>	Avoid negatively worded items whenever possible.
<b>Rule #4:</b>	All of the alternatives should be grammatically consistent with the stem.
<b>Rule #5:</b>	An item should contain only one correct or clearly best answer.
<b>Rule #6:</b>	The relative length of the alternatives should not provide a clue to the correct answer.
<b>Rule #7:</b>	Avoid providing a clue to the answer by using the same or similar words in the stem.
<b>Rule #8</b>	Avoid the use of all-inclusive terms such as never, always, sole, and all in distractors.
<b>Rule #9</b>	Avoid alternatives such as "none of the above" or "all of the above."
<b>Rule #10:</b>	All distractors should be plausible enough to attract trainees who do not know the correct answer.

## **XII. STEPS TO WRITING TEST QUESTIONS**

1. Decide instructional intent - major points
2. Determine what trainee should know - select representative sample of points
3. Write the stem of the item.
4. Write the correct answer.
5. Write 4 or 5 distractors.
6. Check against rules and major principles of test item development.